

A unification of unitary similarity transforms to compressed representations

Raf Vandebril and Gianna M. Del Corso

Abstract

In this paper a new framework for transforming arbitrary matrices to compressed representations is presented. The framework provides a generic way of transforming a matrix via unitary similarity transformations to, e.g., Hessenberg, Hessenberg-like form and combinations of both. The new algorithms are deduced, based on the QR-factorization of the original matrix. Relying on manipulations with rotations, all the algorithms consist of eliminating the correct set of rotations, resulting in a matrix obeying the desired structural constraints. Based on this new reduction procedure we investigate further correspondences such as irreducibility, uniqueness of the reduction procedure and the link with (rational) Krylov methods. The unitary similarity transform to Hessenberg-like form as presented here, differs significantly from the one presented in earlier work. Not only does it use less rotations to obtain the desired structure, also the convergence to rational Ritz-values is not observed in the conventional approach.

Raf Vandebril
Department of Computer Science
KU Leuven, Belgium
Raf.Vandebril@cs.kuleuven.be

Gianna M. Del Corso
Department of Computer Science
Università di Pisa, Italy
DelCorso@di.unipi.it

Article information

- Vandebril, Raf; Del Corso, Gianna M.. *A unification of unitary similarity transforms to compressed representations*, Numerische Mathematik, volume 119, issue 4, pages 641-665, 2011.
- The content of this article is identical to the content of the published paper, but without the final typesetting by the publisher.
- Journal's homepage: <http://link.springer.com/journal/211>
- Published version: <http://dx.doi.org/10.1007/s00211-011-0400-5>
- KU Leuven's repository url: <https://lirias.kuleuven.be/handle/123456789/280650>

A unification of unitary similarity transforms to compressed representations

Raf Vandebril · Gianna M. Del Corso

Received: date / Accepted: date

Abstract In this paper a new framework for transforming arbitrary matrices to compressed representations is presented. The framework provides a generic way of transforming a matrix via unitary similarity transformations to, e.g., Hessenberg, Hessenberg-like form and combinations of both. The new algorithms are deduced, based on the QR -factorization of the original matrix. Relying on manipulations with rotations, all the algorithms consist of eliminating the correct set of rotations, resulting in a matrix obeying the desired structural constraints.

Based on this new reduction procedure we investigate further correspondences such as irreducibility, uniqueness of the reduction procedure and the link with (rational) Krylov methods.

The unitary similarity transform to Hessenberg-like form as presented here, differs significantly from the one presented in earlier work. Not only does it use less rotations to obtain the desired structure, also the convergence to rational Ritz-values is not observed in the conventional approach.

Keywords Unitary similarity transforms, $\{p\}$ -band matrices \cdot $\{p\}$ -semiseparable matrices \cdot Krylov subspaces \cdot Krylov matrices \cdot Ritz-values \cdot Rational Ritz-values

The first author has a grant as “Postdoctoraal Onderzoeker” from the Fund for Scientific Research–Flanders (Belgium). This research was also partially supported by the Research Council K.U.Leuven, project OT/11/055 (Spectral properties of (perturbed) normal matrices and their applications), and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization).

The research of the second author is partially supported by the PRIN project “Analisi di strutture di matrici: aspetti teorici, computazionali e applicazioni” Prot. n. 20083KLJEZ by the Italian MIUR.

R. Vandebril
K.U.Leuven, Dept. of Computer Science, 3001 Leuven(Heverlee), Belgium.
E-mail: raf.vandebril@cs.kuleuven.be

G. M. Del Corso
Università di Pisa, Dept. of Computer Science, Largo Pontecorvo, 3, 56127 Pisa, Italy.
E-mail: delcorso@di.unipi.it

1 Introduction

Eigenvalue computations are mostly composed of two parts. First a unitary similarity transformation to Hessenberg or tridiagonal form [11, 14] is executed. This preprocessing transformation is crucial for the development of efficient and competitive *QR*-algorithms [19, 29, 30]. In some cases, however, the Hessenberg structure is not the most recommended structure and its inverse, the Hessenberg-like matrix can result in faster computations [26] due to a significantly changed convergence behavior.

The first constructive similarity transformations to alternative matrices, admitting low-cost storage, such as semiseparable (plus diagonal) and Hessenberg-like were proposed in [3, 23, 28]. Though interesting convergence results were observed [27], the methods were not competitive due to a nonneglectable extra computational cost w.r.t. the Hessenberg and tridiagonal reduction algorithms.

Detailed studies of the outcome of these reduction algorithms reveal close connections with orthogonal rational functions [6, 21, 22] and relations with rational Krylov sequences [2, 3, 13, 18]. Whereas standard Krylov methods typically have a Ritz-value convergence towards the extreme eigenvalues, rational Krylov techniques are able to shift convergence to other, more interesting points in the plane, by selecting well-chosen poles. This flexibility in altering the convergence behavior is impactful in several applications [20]. The links between Hessenberg-like matrices and rational Krylov spaces enable the development of fast structure exploiting algorithms for retrieving rational Ritz-values. Unfortunately, the first constructive algorithms [3, 23, 28] for retrieving Hessenberg-like (plus diagonal) matrices are not optimal in terms of the number of similarity transformations, losing thereby the valuable link with rational Krylov. Even though theoretical results supporting the existence of such algorithms are available, no constructive methods are provided up to now.

The algorithms proposed in this article are relying on the *QR*-factorization of the involved matrices. Even though using the *QR*-factorization for Hessenberg matrices seems redundant, for many classes of structured rank matrices it provides means for a compact representation. Structured rank matrices are generally dense, and storing the low rank relations can be done effectively for many matrices by storing the *QR*-factorization [8, 25]. An important example is the companion matrix, which is a structured rank Hessenberg matrix. Storing it by a specific *QR*-decomposition enables the development of fast *QR*-algorithms, gaining one order in computation time for the global eigenvalue computations [5, 24].

In this paper, we will first present a new similarity transformation to Hessenberg-like form. Comparing this new approach with the classical methods from [3, 23, 28] we note that this method is much faster (comparable in time to the reduction to Hessenberg form) and moreover it provides the missing link with rational Krylov. The use of the *QR*-factorization for deducing the similarity transformation has several extra advantages. First, when considering both the reduction to Hessenberg and Hessenberg-like form from the *QR*-viewpoint, many similarities become apparent providing a unified framework for both matrix types. Second, starting from the reduction to Hessenberg form, the reduction to generalized Hessenberg form (having more subdiagonals) is trivial; adapting the classical reduction to Hessenberg-like form for retrieving a generalized Hessenberg-like is far from trivial and computationally ex-

uberant expensive. The unifying framework, however, enables us to derive such reductions in a straightforward manner. Finally, utilizing the QR -factorization enables us to derive new reduction algorithms. Though this is more of a theoretical interest it plainly illustrates the new insights obtained by working with the QR -factorization.

The article is organized as follows. Section 2 discusses preliminary results, essential for understanding the article. In Section 3 the unitary similarity transformations to the different structures are presented. In Section 4 it is shown that the reduction process to Hessenberg form or to Hessenberg-like form can be captured in a unifying framework. Section 5 contains the description of the reduction to mixed structures, that is the sum of a generalized Hessenberg and a generalized Hessenberg-like matrix. In Section 6 uniqueness of the reductions is studied. The relation with rational Ritz-values is investigated in Section 7 where it is shown that the reduction to Hessenberg-like form inherits the convergence to the rational Ritz-values rather than the standard convergence. This is illustrated by a numerical experiment where the (rational) Ritz-values are plotted with respect to the order of the submatrix already in Hessenberg(-like) form, revealing that for the reduction to Hessenberg-like form the eigenvalues of this submatrix show a rational Ritz-value convergence behavior.

2 Preliminary results

This section discusses definitions of the involved matrices; graphical schemes necessary for simplifying the understanding of the algorithms; manipulations between rotations and examples underpinning these statements. With $A(i : j, \ell : k)$ we refer to the submatrix of A having columns ℓ to k and rows i to j .

2.1 Definitions

A $\{p\}$ -Hessenberg matrix is a generalization of a standard Hessenberg matrix having all entries below the p -th subdiagonal zero.

Definition 1 The matrix $H = (h_{ij})_{ij} \in \mathbb{C}^{n \times n}$ is a $\{p\}$ -Hessenberg matrix, $p \geq 0$, when $h_{ij} = 0$ for all $i > j + p$. Hence, below the p -th subdiagonal, the matrix is zero.

For symmetric matrices, one obtains a band matrix with bandwidth $2p + 1$. The inverses of $\{p\}$ -Hessenberg matrices are named $\{p\}$ -Hessenberg-like matrices [25] and are of structured rank form.

Definition 2 The matrix Z is called a $\{p\}$ -Hessenberg-like matrix, $p \geq 0$, when $\text{rank}(Z(i : n, 1 : \min\{i + p - 1, n\})) \leq p$ for all valid i . This means that all submatrices taken out of the part below the p -th-superdiagonal have rank at most p .

These matrices are also often referred to as $\{p\}$ -lower semiseparable. Note that Definitions 1 and 2 coincide for $p = 0$ resulting in an upper triangular matrix.

In this paper we will consider also unitary transformations to matrices having a mixed structure, which can, e.g., be expressed as the sum of a $\{p\}$ -Hessenberg and a $\{q\}$ -Hessenberg-like matrix. We will refer to $\{p\}$ -Hessenberg or $\{p\}$ -Hessenberg-like matrices also as generalized Hessenberg or generalized Hessenberg-like matrices, respectively.

2.2 Graphical representations

The forthcoming algorithms are explained by graphical schemes representing the interactions between different rotations and the matrix itself. The algorithms are described by operating on the QR -factorization of the original matrix; this enables a consistent description of the diverse reduction schemes. We introduce the graphical representation by constructing the QR -factorization of a small matrix $A \in \mathbb{C}^{5 \times 5}$. The elements of the matrix $A = (a_{ij})$ are denoted by \times . To construct the QR -factorization, first a_{51} is annihilated by a rotation G_{51}^H . As a result we get $G_{51}^H A$ having element $(5, 1)$ zero. Graphically this is depicted as follows:

$$G_{51}^H A = \begin{array}{c|ccccc} \textcircled{1} & \times & \times & \times & \times & \times \\ \textcircled{2} & \times & \times & \times & \times & \times \\ \textcircled{3} & \times & \times & \times & \times & \times \\ \textcircled{4} & \times & \times & \times & \times & \times \\ \textcircled{5} & \times & \times & \times & \times & \times \\ \hline & 1 & & & & \end{array} = \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}.$$

The bracket with arrows depicts the rotation G_{51}^H acting on row 4 and row 5. The horizontal axis presents a sort of timeline, depicting which rotation needs to be performed first and so forth. The vertical axis numbers the rows. Undetermined elements in the matrices are marked by the symbol \times .

One continues the procedure by successively eliminating the elements a_{41} , a_{31} and finally a_{21} . This results in $G_{21}^H G_{31}^H G_{41}^H G_{51}^H A = Q_1^H A$, having the entire first column except the first element zero. Take $Q_1^H = G_{21}^H G_{31}^H G_{41}^H G_{51}^H$ as the unitary transformation creating zeros throughout the entire first column. Schematically we get the following:

$$\begin{array}{c|ccccc} \textcircled{1} & \times & \times & \times & \times & \times \\ \textcircled{2} & \times & \times & \times & \times & \times \\ \textcircled{3} & \times & \times & \times & \times & \times \\ \textcircled{4} & \times & \times & \times & \times & \times \\ \textcircled{5} & \times & \times & \times & \times & \times \\ \hline & 4 & 3 & 2 & 1 & \end{array} = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}.$$

Consecutively eliminating the elements in the second column below the subdiagonal gives $G_{32}^H G_{42}^H G_{52}^H Q_1^H A = Q_2^H Q_1^H A$:

$$\begin{array}{c|ccccc} \textcircled{1} & \times & \times & \times & \times & \times \\ \textcircled{2} & \times & \times & \times & \times & \times \\ \textcircled{3} & \times & \times & \times & \times & \times \\ \textcircled{4} & \times & \times & \times & \times & \times \\ \textcircled{5} & \times & \times & \times & \times & \times \\ \hline & 7 & 6 & 5 & 4 & 3 & 2 & 1 & \end{array} = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{bmatrix}.$$

Continuing this procedure leads to $Q_4^H Q_3^H Q_2^H Q_1^H A = Q^H A = R$:

$$\begin{array}{c|cccccc} \textcircled{1} & \times & \times & \times & \times & \times \\ \textcircled{2} & \times & \times & \times & \times & \times \\ \textcircled{3} & \times & \times & \times & \times & \times \\ \textcircled{4} & \times & \times & \times & \times & \times \\ \textcircled{5} & \times & \times & \times & \times & \times \\ \hline & 10 & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & \end{array} = \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & 0 & \times & \times \\ 0 & 0 & 0 & 0 & \times \end{bmatrix}. \quad (1)$$

The outcome is the QR -factorization $A = QR$, schematically of the following form:

$$\begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 7 & 6 & 5 & 4 & 3 & 2 & 1 &
 \end{array} \cdot \quad (2)$$

Comparing the representation of the unitary matrix Q^H in Equation (1) and Q in Equation (2), we see that they are represented in a condensed form in (2). The brackets clearly denote on which rows the rotations act, hence some of the rotations commute and their order of performance can be changed. In (2), e.g., two rotations can be executed simultaneously in the third step, one acting on rows 2 and 3 and the other on rows 4 and 5. These schemes contain significant information, since they visibly indicate the order of the transformations and the rows they act on. To compute the QR -factorization of an arbitrary matrix A , $n(n-1)/2$ rotations are required.

Remark 3 In this manuscript we work with rotations, i.e., the determinant equals 1. In fact all results presented here also hold for 2×2 unitary matrices. See [4] for information on rotations (Givens transformations) and how to compute them reliably.

Let us consider the representations of some typical examples such as a Hessenberg, a Hessenberg-like matrix and combinations.

Example 4 (A Hessenberg matrix) Suppose H to be of Hessenberg form with $n = 5$. The QR -factorization (constructed with rotations) $H = QR$ is schematically represented as follows:

$$\begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 4 & 3 & 2 & 1 &
 \end{array} \cdot \quad (3)$$

The matrix Q consists of a *descending* sequence of four rotations. Performing these transformations on the right upper triangular matrix R will fill up the subdiagonal elements and create a Hessenberg matrix.

Example 5 (A Hessenberg-like matrix) Suppose Z to be of Hessenberg-like form. Because of the low rank structure of this matrix, computing the QR -factorization only involves $n-1$ rotations forming an *ascending* sequence. Schematically $Z = QR$ (for $n = 5$) is depicted as follows:

$$\begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 4 & 3 & 2 & 1 &
 \end{array} \cdot \quad (4)$$

Executing the rotations on the upper triangular matrix fills it up with a low rank part. All matrices taken out of the lower triangular part are of rank at most one.

Example 6 (A {2}-Hessenberg-like matrix) Assume $n = 5$ and $p = 2$. Schematically $Z = QR$ is of the following form

$$\begin{array}{c|cccccc} \textcircled{1} & & & & & & \times & \times & \times & \times & \times \\ \textcircled{2} & & & & & & & \times & \times & \times & \times \\ \textcircled{3} & & & & & & & & \times & \times & \times \\ \textcircled{4} & & & & & & & & & \times & \times \\ \textcircled{5} & & & & & & & & & & \times \\ \hline & 6 & 5 & 4 & 3 & 2 & 1 & & & & \end{array},$$

where Q consists of two *ascending* sequences of rotations. Performing the rotations on the upper triangular matrix fills it up with a low rank part, such that all matrices taken out of the part including the superdiagonal have rank at most 2.

Example 7 (Sum of a Hessenberg and Hessenberg-like) Let A be the sum of Hessenberg and an Hessenberg-like matrix. This means that all the submatrices taken out of the part below the subdiagonal have rank at most one. For a 6×6 matrix we have the following situation

$$A = \begin{bmatrix} \times & \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times & \times \\ \boxtimes & \times & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \times & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \times & \times & \times \\ \boxtimes & \boxtimes & \boxtimes & \boxtimes & \times & \times \end{bmatrix},$$

where the entries denoted by \boxtimes represent the rank-one part. Because of the low rank structure, computing the factorization requires only $3(n-2)$ rotations:

$$\begin{array}{c|cccccc} \textcircled{1} & & & & & & \times & \times & \times & \times & \times \\ \textcircled{2} & & & & & & & \times & \times & \times & \times \\ \textcircled{3} & & & & & & & & \times & \times & \times \\ \textcircled{4} & & & & & & & & & \times & \times \\ \textcircled{5} & & & & & & & & & & \times \\ \textcircled{6} & & & & & & & & & & \times \\ \hline & 9 & 8 & 7 & 6 & 5 & 4 & 3 & 2 & 1 & \end{array}. \quad (5)$$

Example 8 (Generic summation) In the general case, where A is the sum of a $\{p\}$ -Hessenberg matrix and a $\{q\}$ -Hessenberg-like matrix, we need $q(2(n-p)-q-1)/2$ rotations to remove the rank- q part. Once we have annihilated the rank- q part, q new diagonals have been formed and we need $\sum_{i=1}^{p+q} (n-i) - 1 = (p+q)(2n-(p+q)-1)/2 - 1$ more rotations to obtain an upper triangular matrix R .

2.3 Manipulating rotations

The forthcoming algorithms depend heavily on manipulating rotations. These operations are already described and proved elsewhere (see e.g. [25,26]). For completeness, however, they are reconsidered without proofs.

Lemma 9 *Let G_1 and G_2 be two rotations, then $G_1 G_2 = G_3$ is again a rotation. We will call this the “fusion” of rotations in the remainder of the text.*

Graphically we depict this as follows:

$$\begin{array}{c|c} \textcircled{1} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \textcircled{2} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \hline & \begin{array}{c} 2 \quad 1 \end{array} \end{array} \text{ resulting in } \begin{array}{c|c} \textcircled{1} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \textcircled{2} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \hline & \begin{array}{c} 1 \end{array} \end{array}.$$

When considering rotations of higher dimensions (say n), we mean that a 2×2 rotation is embedded in the identity matrix of order n . The next lemma states that we can change the order of the rotations.

Lemma 10 (Shift-through operation) *Let 3×3 rotations \check{G}_1, \check{G}_2 and \check{G}_3 be given, such that the rotations \check{G}_1 and \check{G}_3 act on the first two rows of a matrix, and \check{G}_2 acts on the second and third row (when applied on the left to a matrix). Then there exist three rotations \hat{G}_1, \hat{G}_2 and \hat{G}_3 such that*

$$\check{G}_1 \check{G}_2 \check{G}_3 = \hat{G}_1 \hat{G}_2 \hat{G}_3,$$

with \hat{G}_1 and \hat{G}_3 acting on the second and third row and \hat{G}_2 acting on the first two rows.

This result is well-known. The proof is based on the different ways to factor a 3×3 unitary matrix [25]. Schematically we get:

$$\begin{array}{c|c} \textcircled{1} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \textcircled{2} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \textcircled{3} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \hline & \begin{array}{c} 3 \quad 2 \quad 1 \end{array} \end{array} \text{ resulting in } \begin{array}{c|c} \textcircled{1} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \textcircled{2} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \textcircled{3} & \begin{array}{c} \downarrow \\ \downarrow \end{array} \\ \hline & \begin{array}{c} 3 \quad 2 \quad 1 \end{array} \end{array}.$$

The other direction (from the right to the left scheme is depicted by \curvearrowright). We remark that the fusion of rotations is a special instance of the shift-through operation. A final important operation is the shift-through operation of length ℓ .

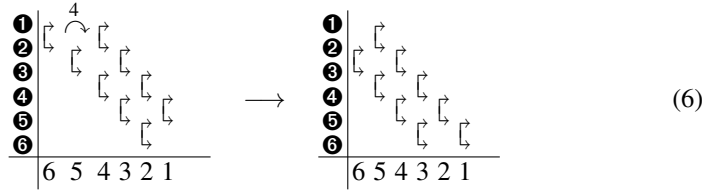
Lemma 11 (Shift-through operation of length ℓ) *Suppose we have the following matrix product GWX , in which G denotes a rotation acting on row 1 and 2. The matrices W and X are both unitary matrices consisting of a descending sequence of ℓ rotations. This means that both W and X consist of ℓ successive rotations. The i th rotation G_i^W of W acts on row $i+1$ and $i+2$. The i th rotation G_i^X of X acts on row i and $i+1$. The matrix product GWX can then be rewritten as*

$$GWX = \hat{W} \hat{X} \hat{G},$$

where \hat{G} is now a rotation acting on row $\ell+1$ and $\ell+2$. The unitary matrices \hat{W} and \hat{X} are again descending sequences of ℓ rotations.

Schematically we obtain that the following two schemes represent the same action of rotations. Hence this represents the same unitary transformation, but factored differently as a sequence of rotations. The sequence W ranges from position 5 to 2, the sequence X ranges from position 4 to 1 and G can be found in position 6. A

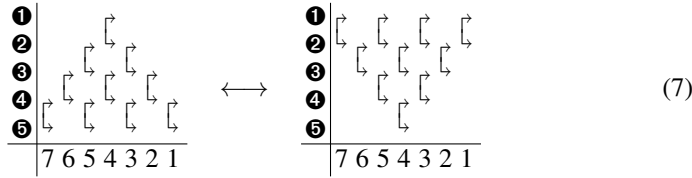
shift-through operation of a specified length is depicted by adding a super or subscript indicating the number of successive shift-through operations that need to be performed.



On the right we see the resulting scheme in which the new sequence \hat{W} ranges from 6 to 3, \hat{X} ranges from 5 to 2 and \hat{G} is located in the first position. To indicate where the marked rotation is going to and to specify the desired number of single shift-through to complete the task, we use the symbols $\overset{\ell}{\curvearrowright}$, $\overset{\ell}{\curvearrowleft}$, $\overset{\ell}{\curvearrowright}$ and $\overset{\ell}{\curvearrowleft}$.

Example 12 (Factorization of a unitary matrix) An interesting result of the shift-through lemma is the factorization of a unitary $n \times n$ matrix. The shift-through lemma illustrates already a sort of \vee and \wedge -pattern for factoring a 3×3 unitary matrix.

An arbitrary unitary matrix U can be factored in sequences of rotations as the left matrix in Scheme (7), in fact one just computes the QR -factorization and R becomes the identity matrix¹. Applying successive shift-through operations can change the \wedge -form on the left in a \vee -form on the right.



This example states in a certain sense that the QR -factorization of an arbitrary matrix can also be computed alternatively, more details can be found in [25].

3 The unitary similarity transforms

The algorithms deduced in this section work on the QR -factorization of the given matrix A and unitarily transform this QR -factorization to the desired representation.

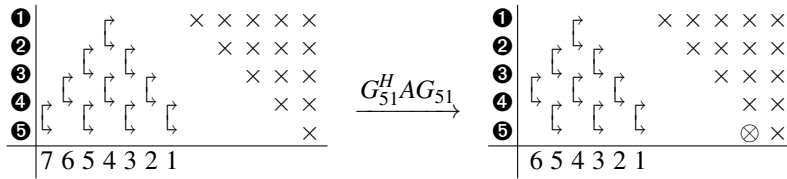
3.1 Unitary similarity transformation to Hessenberg form

The unitary similarity transformation of an arbitrary matrix to Hessenberg form is well-known [14, 29]. Let us alternatively interpret this reduction starting from the QR -factorization of the initial matrix. The description here appears to be different from the standard reduction technique, one can prove, however, (see Section 6) that both approaches are identical.

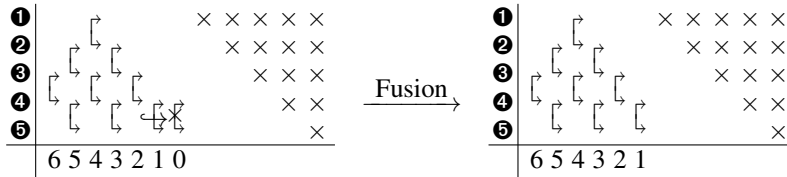
¹ The matrix R equals the identity except possibly the trailing element $r_{nn} = \det U$. The appearance of this unimodular factor does, however, not pose any difficulty.

Assume a dense matrix is given having a QR -factorization of the form (2) (rotations are allowed to equal the identity, this does not pose any problems). Assume that A has dimensions 5×5 , the reader can infer how to deal with the general case. We have $A = QR = Q_1 Q_2 Q_3 Q_4 R$, where Q_1^H annihilates elements in the first column of A , Q_2^H in the second column and so on. Each matrix Q_i is a combination of rotations $Q_1 = G_{51} G_{41} G_{31} G_{21}$, where G_{51}^H annihilates element a_{51} and so forth.

In the end we would like to obtain a factorization as in Scheme (3). To achieve this goal all rotations in the QR -factorization, except $G_{21}, G_{32}, G_{43}, G_{54}$ need to be removed. Let us start by removing one rotation at a time by the designated unitary similarity transformations. The rotation G_{51} is expunged by a similarity transformation determined by G_{51} itself. We get $A^{(51)} = G_{51}^H A G_{51} = G_{41} G_{31} G_{21} Q_2 Q_3 Q_4 R G_{51}$. Schematically we get the following transition, in the right scheme R is transformed into $R G_{51}$ having a bulge in position $(5, 4)$.

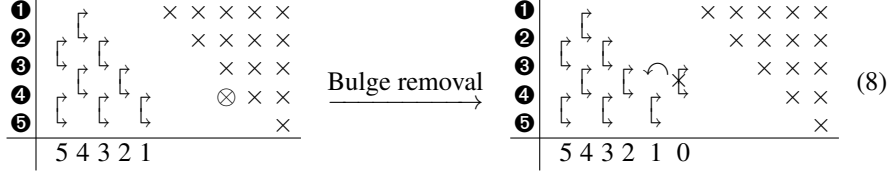


To obtain the QR -factorization of a Hessenberg matrix we need to remove the bulge in the R -factor in position $(5, 4)$. A rotation \tilde{G} is constructed such that $R^{(51)} = \tilde{G}(R G_{51})$ is upper triangular again $A^{(51)} = G_{41} G_{31} G_{21} Q_2 Q_3 Q_4 \tilde{G}^H \tilde{G} R G_{51} = G_{41} G_{31} G_{21} Q_2 Q_3 Q_4 \tilde{G}^H R^{(51)}$. We can remove \tilde{G}^H by a fusion with the lower right rotation of Q_4 . We remark that in the previous scheme, the left scheme represented A and the right scheme represented $A^{(51)}$. The schemes below represent twice $A^{(51)}$, with a slightly changed representation. On the left \tilde{G}^H is marked with a \times symbol since it is undesired. Even though it is not visibly depicted in the scheme on the right, the transformation in position 1 has changed because of the fusion.

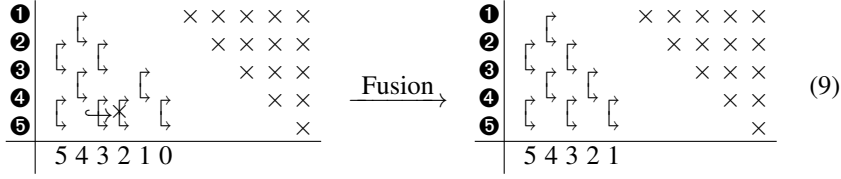


Since the fusion altered a rotation in the unitary factor Q_4 we have the following factorization for $A^{(51)} = G_{41} G_{31} G_{21} Q_2 Q_3 Q_4^{(51)} R^{(51)}$. The next unitary similarity is determined by G_{41} and removes this rotation from the factorization. Performing it onto $A^{(51)}$, will create a new bulge in the matrix $R^{(51)}$. The matrix $A^{(41)} = G_{41} A^{(51)} G_{41}^H$ is depicted in the left of Scheme (8). On the right we see that another rotation \tilde{G} is

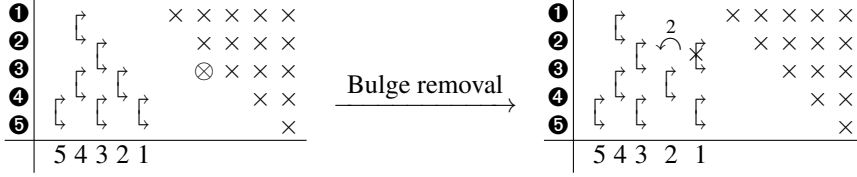
used for annihilating the bulge.



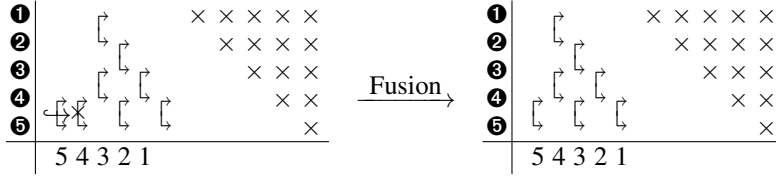
Unfortunately, a fusion cannot be applied this time, first a shift-through operation is needed. After the shift-through operation we get the left of Scheme (9) in which a new fusion is depicted and the rotation disturbing the nice structure is again marked with \times . After a fusion we obtain the right figure, which represents the QR -factorization of the matrix $A^{(41)} = G_{31}G_{21}Q_2Q_3^{(41)}Q_4^{(41)}R^{(41)}$.



Removal of G_{31} is carried out by a similarity transformation $A^{(31)} = G_{31}^H A^{(41)} G_{31} = G_{21}Q_2Q_3^{(41)}Q_4^{(41)}R^{(41)}G_{31}$. The matrix $A^{(31)}$ has a bulge its R factor (position (3, 2)).



The shift-through operation of length 2 is depicted. To finish a fusion is performed.



We have now a factorization of the form $A^{(31)} = G_{21}Q_2^{(31)}Q_3^{(31)}Q_4^{(31)}R^{(31)}$. We will not remove the transformation G_{21} , since it is a part of the QR -factorization of the Hessenberg matrix.

Next one removes all but one rotation from the unitary matrix $Q_2^{(31)}$. On the left of Scheme (10) these rotations are marked with a \times . Similar techniques as the ones

discussed before can be used to retrieve the right scheme, where the final rotation designated for removal is marked by a \times .

$$\begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 5 & 4 & 3 & 2 & 1 & & &
 \end{array}
 \xrightarrow{\text{Similarity}}
 \begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 5 & 4 & 3 & 2 & 1 & & &
 \end{array}
 \quad (10)$$

After removing this final rotation in position 3, a Hessenberg matrix is obtained.

3.2 Unitary similarity transformation to Hessenberg-like form

The reduction to Hessenberg-like form proceeds similar to the reduction to Hessenberg form. Whereas in the reduction to Hessenberg form rotations were removed from the left side of the Q -factor, now they will be removed from the right side. In the end, all rotations except G_{51}, G_{41}, G_{31} and G_{21} are removed. The startup phase slightly differs as the rotations designated for removal have to be brought to the right of the R -factor. Let $A = Q_1 Q_2 Q_3 G_{54} R$, first we apply G_{54} to R as can be seen in Scheme (11)

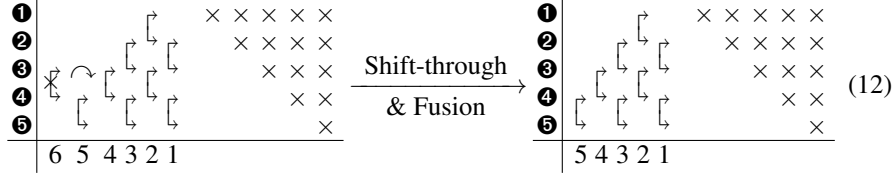
$$\begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 6 & 5 & 4 & 3 & 2 & 1 & &
 \end{array}
 \quad (11)$$

The matrix $G_{54}R$ has a bulge in position (5, 4). The rotation \tilde{G}_{54} determines the similarity transformation and is designed such that $(G_{54}R)\tilde{G}_{54} = R^{(54)}$ is upper triangular again. Hence, we get $A^{(54)} = \tilde{G}_{54}^H A G_{54} = \tilde{G}_{54}^H Q_1 Q_2 Q_3 R^{(54)} = Q_1^{(54)} Q_2 Q_3 R^{(54)}$. The rotation \tilde{G}_{54}^H is removed by a fusion $\tilde{G}_{54}^H Q_1 = Q_1^{(54)}$ as depicted in the left scheme and executed in the right scheme.

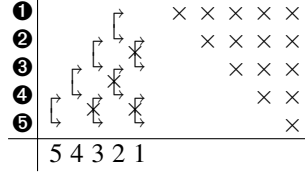
$$\begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 7 & 6 & 5 & 4 & 3 & 2 & 1 &
 \end{array}
 \xrightarrow{\text{Fusion}}
 \begin{array}{c|cccccc}
 \textcircled{1} & & & & \times & \times & \times & \times & \times \\
 \textcircled{2} & & & & & \times & \times & \times & \times \\
 \textcircled{3} & & & & & & \times & \times & \times \\
 \textcircled{4} & & & & & & & \times & \times \\
 \textcircled{5} & & & & & & & & \times \\
 \hline
 & 6 & 5 & 4 & 3 & 2 & 1 & &
 \end{array}$$

The global flow of the method should be clear now. In the next step one removes G_{43} from $A^{(54)} = Q_1^{(54)} Q_2 G_{53} G_{43} R^{(54)}$. This task is accomplished by a similarity transformation on $A^{(54)}$ with rotation \tilde{G}_{43} determined such that $(G_{43}R^{(54)})\tilde{G}_{43} = R^{(43)}$ is upper triangular. On the left of Scheme (12) the matrix $A^{(43)} = \tilde{G}_{43}^H A^{(54)} \tilde{G}_{43}$ is

shown, with a shift-through operation depicted. The shift-through operation is followed by a fusion, absorbing the redundant rotation, to obtain the right scheme.



After removing the marked rotations from the next scheme, the QR -factorization of a Hessenberg-like matrix is obtained.



Remark 13 Reducing a matrix to Hessenberg form by the classical procedure uses $(n-1)(n-2)/2$ similarity transformations, each determined by a single rotation. The original reduction to Hessenberg-like form as proposed in [23] uses twice as many similarity transformations (each transformation determined by a single rotation).

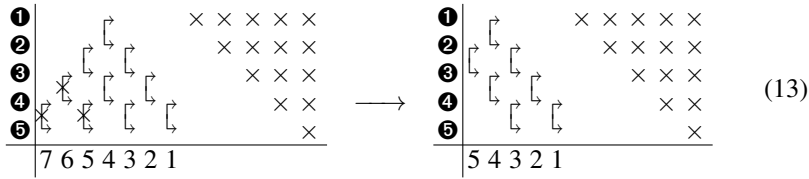
This novel technique is, however, capable of achieving the Hessenberg-like form by only $(n-1)(n-2)/2$ unitary similarity transformations. This significantly different approach is therefore computationally advantageous w.r.t. the classical way.

3.3 Unitary similarity transformation to generalized Hessenberg form

The new algorithm is not only computationally cheaper (see Remark 13) but also easily extends to reductions to $\{p\}$ -Hessenberg-like form as opposed to the classical approach [23]. Let us first study the reduction to $\{p\}$ -Hessenberg form.

The Q -factor in the QR -factorization of a standard Hessenberg matrix consists of a single sequence of descending rotations. In case of a $\{p\}$ -Hessenberg matrix, we will have p descending sequences of rotations. For example, on the right of Scheme (13), the structure of the QR -factorization of a $\{2\}$ -Hessenberg matrix is visualized.

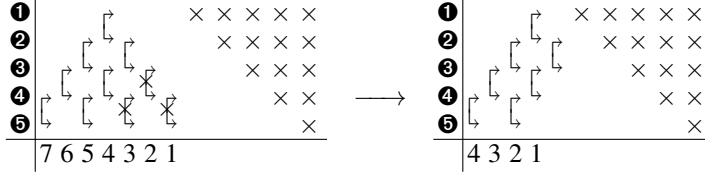
The same procedure as for the reduction to Hessenberg form is followed. The only difference is the removal of less rotations. In the QR -factorization (13) the rotations marked with \times should be removed to obtain a $\{2\}$ -Hessenberg matrix.



It is obvious that for retrieving a $\{p\}$ -Hessenberg matrix, only $(n-p)(n-p-1)/2$ similarity transformations with rotations need to be performed.

3.4 Unitary similarity transformation to generalized Hessenberg-like form

For a Hessenberg-like matrix, the Q -factor in the QR -factorization consists of one ascending sequence of rotations. In the $\{p\}$ -Hessenberg-like case this will be p sequences of ascending rotations. Schematically, for the $p = 2$ case we get the right figure of Scheme (3.4). Removal of the rotations marked in Scheme 3.4 results in a $\{2\}$ -Hessenberg-like matrix.

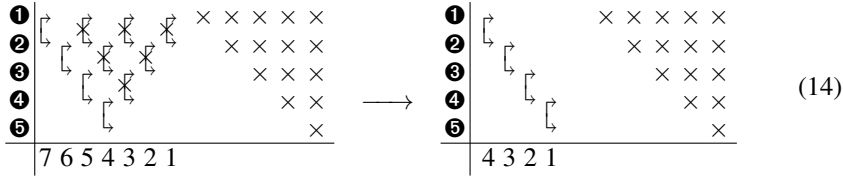


Remark 14 Obviously, reducing a matrix to $\{p\}$ -Hessenberg-like form is a simplification of the transformation to Hessenberg-like form. This is a big difference with the technique from [23], which is unable to attain the generalized form.

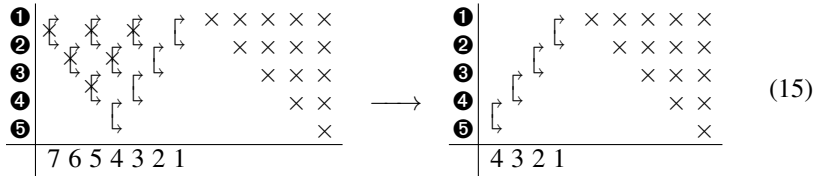
4 Inner, outer and other annihilation schemes

Reconsidering the previous section, one can conclude that transforming matrices to (generalized) Hessenberg form is a kind of outer annihilation process, removing thereby the rotations on the exterior of the QR -factorization. Similarly the reduction to (generalized) Hessenberg-like form can be considered as an inner annihilation process where the rotations in the interior of the QR -factorization are designated for removal. However, whether the rotations to be removed are positioned in the interior or the exterior depends entirely on the rotational factorization of the Q -factor. Example 12 revealed different ways for factoring a unitary matrix. Consider now, for example, a QR -factorization where the matrix Q is in the \vee - instead of the \wedge -form.

Based on the \vee -pattern, the reduction to (generalized) Hessenberg form becomes an inner annihilation process as illustrated in Scheme (14).

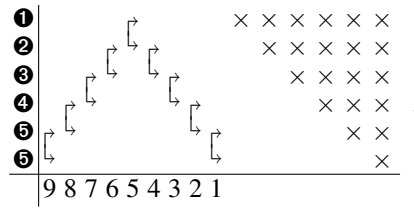


Similarly, the reduction to (generalized) Hessenberg-like form becomes an outer annihilation process as shown in Scheme (15).



In [9, 10] it was investigated how to construct the QR -factorization minimizing thereby the number of rotations for representing the Q -factor (an arbitrary Q -factor requires $n(n-1)/2$ rotations). For a dense unstructured matrix the above reduction procedures require the removal of $(n-2)(n-1)/2$ rotations to obtain the Hessenberg or Hessenberg-like form. Evidently, whenever the Q -factor has strictly less than $n(n-1)/2$ rotations in its factorization, this has a positive impact on the complexity of the reduction algorithms and results in an essentially lower number of rotations to be removed. To illustrate the unifying framework of the presented algorithms some additional examples of particular structured rank matrices and their reduction algorithms are given.

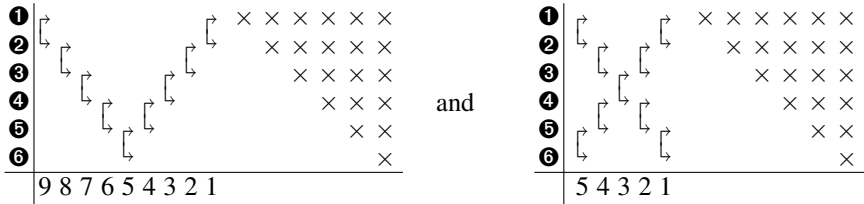
Example 15 (Quasiseparable) As an example, we will propose how to reduce a quasiseparable matrix to Hessenberg form. These algorithms are proposed in [12, 17]. The QR -factorization of a quasiseparable matrix is of the form (\wedge -pattern):



Exploiting techniques from the previous sections it is clear how to reduce the matrix to Hessenberg(-like) form. Removing the left(right) leg of the \wedge is done by removing a rotation and then chasing the perturbation, just like proposed in the articles [12, 17].

A feature of utilizing the QR -factorization for designing particular unitary similarity transformations is that the various factorizations of the unitary matrix Q also leads to other variants of the reduction procedures.

Example 16 (Quasiseparable two-way reduction) Reducing the quasiseparable matrix to Hessenberg form via the \wedge -pattern leads to an outer annihilation scheme. Consider the following variants of the QR -factorization of the quasiseparable matrix:

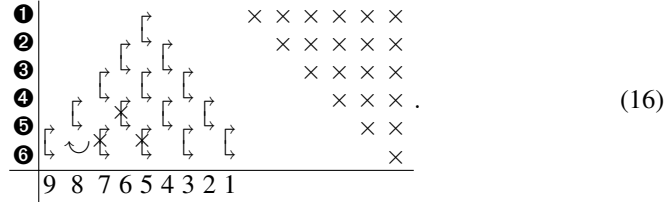


The left scheme (\vee -pattern) results in an inner annihilation process for obtaining a Hessenberg matrix. The right scheme (\times -pattern) is a combination of both an inner and an outer annihilation process. It is interesting that one can start simultaneously by removing the lower left and the upper right leg of the \times resulting in a two-way reduction to Hessenberg form. This can even be implemented in a parallel fashion.

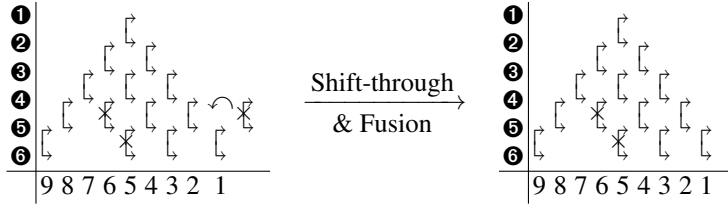
5 Mixed reductions

The reduction to a mixed structure, that is to the sum of a $\{p\}$ -Hessenberg plus a $\{q\}$ -Hessenberg-like matrix is a bit more involved since we have to remove rotations appearing in the middle of the sequences of rotations.

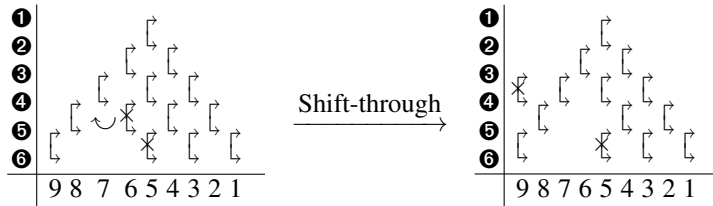
Let us describe the idea following a particular example for reducing a 6×6 matrix A to the sum of a Hessenberg-like and a Hessenberg matrix as appearing in Equation (5). In Scheme (16), the rotations that need to be removed are marked.



We have $A = QR = Q_1 G_{62} G_{52} G_{63} Q_4 Q_5 R$, where we denote by Q_1 the left sequence of ascending rotations, and by Q_4 and Q_5 the two right most sequences of descending rotations. Once the three central rotations are removed by unitary similarity transformations the QR -factorization of a Hessenberg plus Hessenberg-like matrix is obtained. By means of a shift-through operation (the operation is shown in Scheme (16)) one can push G_{62} to the front of the factorization $A = \tilde{G}_{62} Q_1^{(62)} G_{52} G_{63} Q_4 Q_5 R$. Using \tilde{G}_{62} for the unitary similarity transformation on A we get $A^{(62)} = \tilde{G}_{62}^H A \tilde{G}_{62}$, where $R \tilde{G}_{62}$ has a bulge in position (5,4). Recovering the upper triangular structure of R affects the Q -factor as shown in the next scheme.

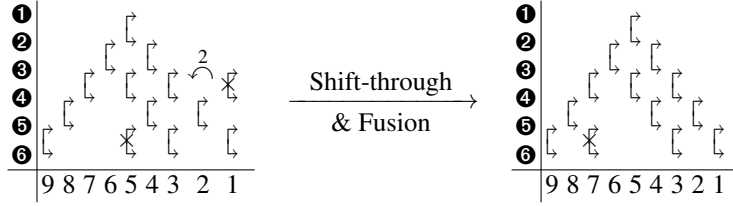


We have $A^{(62)} = Q_1^{(62)} G_{52} G_{63} Q_4^{(62)} Q_5^{(62)} R^{(62)}$. Next we remove G_{52} similarly as G_{63} , the transformation is brought to the front of the Q -factor by a shift-through operation.



\tilde{G}_{52} denotes the leftmost rotation and can be removed by a unitary similarity transformation. We get $A^{(62)} = \tilde{G}_{52} Q_1^{(52)} G_{63} Q_4^{(62)} Q_5^{(62)} R^{(62)}$ and $A^{(52)} = \tilde{G}_{52}^H A^{(62)} \tilde{G}_{52} =$

$Q_1^{(52)} G_{63} Q_4^{(62)} Q_5^{(62)} R^{(52)}$, with $R^{(52)}$ having a bulge in position $(4, 3)$. A new rotation is introduced to remove the bulge in position $(4, 3)$ altering again the Q -factor.



We get $A^{(52)} = Q_1^{(52)} G_{63}^{(52)} Q_4^{(52)} Q_5^{(52)} R^{(52)}$. The procedure to remove the last rotation $G_{63}^{(52)}$ is identical to that performed to remove G_{61} at the beginning, and consists of a shift-through operation to bring a rotation outside, a similarity transformation which produces a bulge in the factor R and a new rotation to remove the bulge, which will be incorporated in the chain $Q_4^{(52)}$. At the end of this procedure we obtain the QR -factorization as in (5).

Remark 17 The key idea for the reduction process is to remove rotations through the fusion operation. Note that the shift-through operation allows to move up or down rotations. In these schemes, where we have series of ascending or descending rotations, in order to apply a fusion we have to bring the undesired rotation to the bottom or top two rows. Hence, if we have a rotation that acts on rows $n-i-1, n-i$, we need a sequence of i descending rotations to move it down accordingly with Lemma 11, and one more rotation acting on rows $n-1, n$ to be fused with. This is the reason why this scheme does not fit the reduction to semiseparable plus diagonal form, or in general the reduction to generalized Hessenberg-like plus diagonal matrices.

6 Uniqueness of the reduction procedure

In this section we will see how the QR -factorization can play an important role in proving, justifying and unifying many results related to Hessenberg and Hessenberg-like matrices. This section includes a unifying treatment of the properness of Hessenberg and Hessenberg-like matrices and discusses uniqueness issues in the outcome of the similarity transformations.

6.1 Properness of Hessenberg and Hessenberg-like matrices

The definitions of properly Hessenberg and properly Hessenberg-like matrices are slightly different from each other. A matrix is said to be properly (or irreducible, or unreduced) Hessenberg if all of its subdiagonal elements are nonzero [14]. A matrix is said to be properly (or unreduced) Hessenberg-like if none of the subdiagonal elements equals zero and if none of the superdiagonal elements is includable in the lower triangular rank structure [26]. These definitions find their origin in the QR -algorithm. They provide easy verifiable criteria such that, e.g., an implicit QR -step can run to

completion and will theoretically not breakdown. The underlying idea is, however, that the Q -factor in the QR -factorization is essentially (up to unitary diagonal scaling) unique. A sufficient requirement is uniqueness of the first $n - 1$ columns of Q , thereby uniquely determining the trailing column of Q . Conveying this sufficient condition to a Hessenberg matrix translates to the straightforward requirement that none of the subdiagonal elements annihilates. Adapting this to the Hessenberg-like case is more involved putting conditions on not only the sub-, but also the superdiagonal elements. Examining, however, the QR -factorization of these matrices we get a unifying definition of properness.

Let H be a proper Hessenberg matrix, the assumption that none of the subdiagonal entries vanish implies that the rotations for bringing the Hessenberg matrix to upper triangular form are essentially uniquely defined and different from the identity. As a result, the diagonal elements of the resulting upper triangular matrix R are nonzero, except possibly the trailing element r_{nn} . As a result we get the following necessary and sufficient condition of properness of a Hessenberg matrix. In Lemma 19 this condition is conveyed to the Hessenberg-like case.

Lemma 18 *A Hessenberg matrix $H \in \mathbb{C}^{n \times n}$, having QR -factorization $H = QR$, is properly Hessenberg if and only if, the rotations appearing in the factorization of the unitary matrix Q differ from the identity and $R(1 : n - 1, 1 : n - 1)$ is nonsingular.*

Lemma 19 *A Hessenberg-like matrix $Z \in \mathbb{C}^{n \times n}$, having QR -factorization $Z = QR$, is properly Hessenberg-like if and only if, the rotations appearing in the factorization of the unitary matrix Q differ from the identity and $R(1 : n - 1, 1 : n - 1)$ is nonsingular.*

It is almost trivial to verify that the conditions in Lemmas 18 and 19 coincide with the classical definitions of properness.

6.2 Unitary similarity transformations

Generic uniqueness results on unitary similarity transformations to Hessenberg(-like) form, not related to a specific algorithm, can be found in [2, 3, 29] and an implicit Q -theorem for Hessenberg(-like) matrices is presented in [3, 14, 26].

In this section we will link the unitary similarity transformations to Hessenberg and Hessenberg-like form to the QR -factorizations of particular (rational) Krylov subspaces. These relations enable a simple alternative proof of uniqueness as in [29].

Denote Krylov subspaces as $\mathcal{K}_k(A, \mathbf{v}) = \text{span}\{\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^{k-1}\mathbf{v}\}$. The relationship between upper Hessenberg matrices and Krylov subspaces is detailed in [29]. In the same way, we can prove a similar relationship for Hessenberg-like matrices and rational Krylov subspaces, that is $\mathcal{K}_k(A^{-1}, \mathbf{v}) = \text{span}\{\mathbf{v}, A^{-1}\mathbf{v}, A^{-2}\mathbf{v}, \dots, A^{-(k-1)}\mathbf{v}\}$. The next theorem states that a unitary matrix whose columns form a basis for a rational Krylov subspace uniquely determine a unitary similarity transformation to Hessenberg-like form (see also [29, Theorem 3.3.2] which is the analog of Theorem 20 for Hessenberg matrices).

Theorem 20 *Let $A \in \mathbb{C}^{n \times n}$ and $Z \in \mathbb{C}^{n \times n}$ be nonsingular matrices, $V \in \mathbb{C}^{n \times n}$ a unitary matrix, such that $Z = V^H A V$. Denote the columns of V by $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.*

a) For Z being a non-proper Hessenberg-like matrix, with k such that $z(i, j) = 0, i \geq k, 1 \leq j < k$, we get

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j\} = \mathcal{K}_j(A^{-1}, \mathbf{v}_1), \quad j = 1, \dots, k.$$

b) For Z being a proper Hessenberg-like, we get

$$\text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j\} = \mathcal{K}_j(A^{-1}, \mathbf{v}_1), \quad j = 1, \dots, n. \quad (17)$$

In this case $\mathcal{K}_n(A^{-1}, \mathbf{v}_1)$ has dimension n .

c) Conversely, if (17) holds, Z is a Hessenberg-like matrix.

Proof The inverse of a Hessenberg-like matrix is of Hessenberg form [25]. In case A is non-proper, its inverse A^{-1} is also non-proper, having the same zero block below the diagonal. As a result we can apply Theorem 3.3.2 of [29] to matrix A^{-1} .

An alternative, direct proof is obtained by using the QR -factorization as follows. Let $A = QR$ and $Z = \hat{Q}\hat{R}$. As \hat{Q} consists of a sequence of ascending rotations, \hat{Q}^H is an upper Hessenberg matrix. Take $\mathbf{v} = \mathbf{v}_1$, we have

$$\begin{aligned} \mathcal{K}_j(A^{-1}, \mathbf{v}) &= \text{span}\{\mathbf{v}, V(\hat{Q}\hat{R})^{-1}V^H\mathbf{v}, V(\hat{Q}\hat{R})^{-2}V^H\mathbf{v}, \dots, V(\hat{Q}\hat{R})^{-j+1}V^H\mathbf{v}\} \\ &= \text{span}\{\mathbf{v}, V(\hat{Q}\hat{R})^{-1}\mathbf{e}_1, V(\hat{Q}\hat{R})^{-2}\mathbf{e}_1, \dots, V(\hat{Q}\hat{R})^{-j+1}\mathbf{e}_1\} \\ &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_j\}. \end{aligned}$$

If Z is a proper Hessenberg-like the equality holds for each $j = 1, \dots, n$, otherwise it holds only up to the first index k such that $\hat{r}_{kk} = 0$.

Remark 21 For a singular A the rational Krylov space is not defined, the unitary similarity transformation, does, however still exist. Infinitesimal perturbations of Z resolve this singularity problem. With the same notation as in Theorem 20, Z proper, but singular, we perturb the trailing element of Z : $\tilde{Z} = Z + \epsilon \mathbf{e}_n \mathbf{e}_n^T$ resulting in a non-singular $\tilde{A} = V \tilde{Z} V^H$. Applying Theorem 20 to \tilde{A} and \tilde{Z} , we get that

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_j\} = \mathcal{K}_j(\tilde{A}^{-1}, \mathbf{v}_1), \quad j = 1, \dots, n.$$

For Z a not proper and singular Hessenberg-like matrix, we have to perturb one or more entries corresponding to the zeros in the R -factor of the QR -factorization of Z . In this case the equality only holds for the first $j \leq k$ columns of V spanning $\mathcal{K}_j(\tilde{A}^{-1}, \mathbf{v})$, for $j = 1, \dots, k$, where k locates the first zero on the diagonal of R .

The following theorem states uniqueness of the reduction procedure.

Theorem 22 (Implicit Q -theorem) Consider a nonsingular matrix A . Let V_1 and V_2 be two unitary matrices sharing the same first column (up to a unimodular factor) such that

$$Q_1 R_1 = Z_1 = V_1^H A V_1 \quad \text{and} \quad Q_2 R_2 = Z_2 = V_2^H A V_2,$$

where the Z_1 and Z_2 are two Hessenberg-like matrices. Let k_1 (resp. k_2) be the position of the first zero on the diagonal of R_1 (resp. R_2), this is a measure of properness. Then, $Z_1(1:k, 1:k)$ and $Z_2(1:k, 1:k)$ are essentially equal, with $k = \min\{k_1, k_2\}$.

Proof The matrix $K_j(A, \mathbf{v})$ has the vectors $\mathbf{v}, A\mathbf{v}, \dots, A^{j-1}\mathbf{v}$ as columns. Using standard properties of Krylov spaces we get (σ is a unimodular factor) :

$$\begin{aligned} V_1 K_j(Z_1^{-1}, \mathbf{e}_1) &= K_j(V_1 Z_1^{-1} V_1^H, V_1 \mathbf{e}_1) = K_j(A^{-1}, V_1 \mathbf{e}_1) \\ &= \sigma K_j(A^{-1}, V_2 \mathbf{e}_1) = \sigma K_j(V_2 Z_2^{-1} V_2^H, V_2 \mathbf{e}_1) = \sigma V_2 K_j(Z_2^{-1}, \mathbf{e}_1), \end{aligned}$$

for $1 \leq j \leq k$. The first three equalities hold for $1 \leq j \leq k_1$, the last three hold for $1 \leq j \leq k_2$, by Theorem 20. The left and the right terms are QR -factorizations. Essential uniqueness of the QR -factorization implies that the first k columns of the matrices V_1 and V_2 are essentially identical proving thereby the Theorem.

The approach, as presented, here is not the only possibility for proving uniqueness. Direct calculations exploiting the intrinsic low rank structure of Hessenberg-like matrices are also possible [26]. Also the link to standard (reverted) Krylov subspaces can be used [2], here the authors exploit the link

$$\text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_j\} = \text{span}\{A^{n-1}\mathbf{v}, \dots, A^{n-j}\mathbf{v}\}, \quad j = 1, \dots, n-1,$$

when considering \mathbf{v} as the solution of the system $A^{n-1}\mathbf{v} = \mathbf{v}_1$.

7 Ritz-values and the new similarity transformations

In this section, we will prove that the new similarity transformation for bringing the matrix A to Hessenberg-like form inherits rational Krylov behavior, i.e., the eigenvalues in that part of the matrix already in Hessenberg-like form are rational Ritz-values. The reduction to Hessenberg form and the classical Hessenberg-like [27] procedure, inherit, on the contrary, the classical Ritz-value convergence behavior.

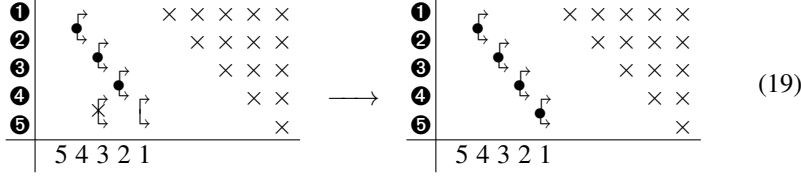
7.1 Part of the matrix in the desired structure

Let us clarify which parts of the matrices during the unitary similarity transformations proposed in this article are already of the desired structure. We will indicate, by graphical schemes, in an interactive fashion, which rotations remain unaltered during the remaining reduction procedure. As they will not alter anymore, these transformations make up the part of the matrix in final form. In the next schemes, the rotations that will not be affected anymore by fusions and shift-throughs are marked with a \bullet .

We commence with the reduction to Hessenberg form. In the first phase of the reduction process three rotations are removed (marked with \times), this corresponds to bringing the first column of the matrix A in Hessenberg form. We note that in the removal process of these three rotations the top transformation will not change. This implies that the upper left element of the matrix a_{11} was already in final Hessenberg form before executing the similarity transformations, hence the \bullet .

$$\begin{array}{c|ccccc} \textcircled{1} & & & & & & \\ \textcircled{2} & & & & & & \\ \textcircled{3} & & & & & & \\ \textcircled{4} & & & & & & \\ \textcircled{5} & & & & & & \\ \hline & 7 & 6 & 5 & 4 & 3 & 2 & 1 \end{array} \longrightarrow \begin{array}{c|ccccc} \textcircled{1} & & & & & & \\ \textcircled{2} & & & & & & \\ \textcircled{3} & & & & & & \\ \textcircled{4} & & & & & & \\ \textcircled{5} & & & & & & \\ \hline & 5 & 4 & 3 & 2 & 1 \end{array} \quad (18)$$

The result is the matrix $A^{(31)}$ depicted in the right of Scheme (18), another rotation is now in final form and will not change when removing the next two rotations, marked by \times . Denote the global reduction to Hessenberg form with $V^H AV = H$. Scheme (18) indicates the following equalities²: for the left scheme we get that $a_{11} = \mathbf{e}_1^T A \mathbf{e}_1 = \mathbf{e}_1^T H \mathbf{e}_1$ and for the right scheme we get that $A^{(31)}(1:2, 1:2) = [\mathbf{e}_1, \mathbf{e}_2]^T A^{(31)} [\mathbf{e}_1, \mathbf{e}_2] = [\mathbf{e}_1, \mathbf{e}_2]^T V^H AV [\mathbf{e}_1, \mathbf{e}_2] = [\mathbf{e}_1, \mathbf{e}_2]^T H [\mathbf{e}_1, \mathbf{e}_2]$ are already of Hessenberg form. More precisely we get $a_{11} = h_{11}$ and $A^{(31)}(1:2, 1:2) = H(1:2, 1:2)$. Removal of two marked rotations in the right of Scheme (18) gives the left of Scheme (19).

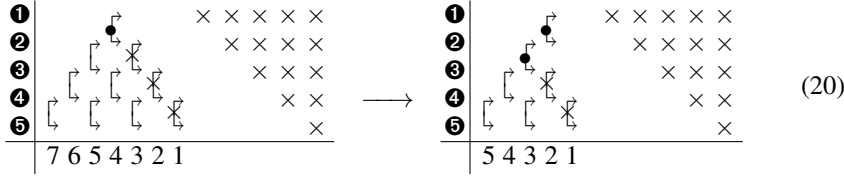


The left of Scheme (19) (the first two columns are in Hessenberg form) is equivalent to $A^{(42)}(1:3, 1:3) = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]^T A^{(42)} [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] = [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]^T H [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3] = H(1:3, 1:3)$ and the right scheme (the first three columns are in Hessenberg form) implies

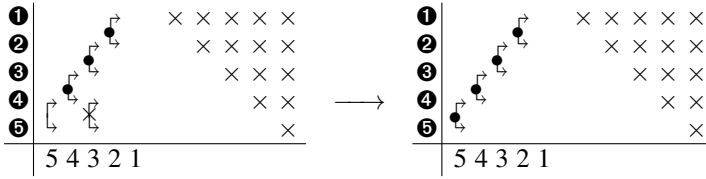
$$\begin{aligned} A^{(53)}(1:4, 1:4) &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4]^T A^{(53)} [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4] \\ &= [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4]^T V^H AV [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3, \mathbf{e}_4] = H(1:4, 1:4). \end{aligned}$$

Since $A^{(53)} = H$, we get full equality here.

For the reduction to Hessenberg-like form, we get similar results. Let us illustrate compactly which parts of the matrix will reach final form after a number of similarity transformations. Assume in the end $V^H AV = Z$ of Hessenberg-like form.



When starting the reduction procedure (left of Scheme (20)) the upper left part of A is already in final form, implying that $\mathbf{e}_1^T V^H AV \mathbf{e}_1$ is fixed throughout the remainder of the procedure. In the right of Scheme (20) we note that the upper left 2×2 subblock of the matrix corresponding to $[\mathbf{e}_1, \mathbf{e}_2]^T V^H AV [\mathbf{e}_1, \mathbf{e}_2]$ is of the desired Hessenberg-like structure and will not alter anymore. In the following two schemes clearly the part that will remain unaltered grows.



² We assume $V\mathbf{e}_1 = \mathbf{e}_1$. Another initial similarity transformation to obtain a generic $V\mathbf{e}_1$ is possible and does not complicate matters. It coincides with a different starting vector for obtaining the Ritz-values.

7.2 Relation with the (rational) Ritz-values

The Ritz-values of a matrix A are defined as the eigenvalues of the projected counterpart $V^H A V$, where the columns of $V = [\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k]$ form an orthogonal basis for the Krylov subspace $\mathcal{K}_k(A, \mathbf{v}_1) = \text{span}\{\mathbf{v}_1, A\mathbf{v}_1, A^2\mathbf{v}_1, \dots, A^{k-1}\mathbf{v}_1\}$, i.e. that $\text{span}\{\mathbf{v}_1, A\mathbf{v}_1, A^2\mathbf{v}_1, \dots, A^{k-1}\mathbf{v}_1\} = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_k\}$. For rational Ritz-values (without shifts) we consider the Krylov subspace $\mathcal{K}_k(A^{-1}, \mathbf{v})$. In [1, 7, 15, 16] the convergence of (rational) Ritz-values to eigenvalues is studied and it is proven that under mild conditions on, e.g., the starting vector, the Ritz-values approximate those eigenvalues well-separated from the rest of the spectrum first. In some circumstances these are not the eigenvalues one wants to find first, as a result shift and invert techniques are used resulting in rational Ritz-values.

When transforming a matrix to Hessenberg form, it is known that the part of the matrix already in Hessenberg form has as eigenvalues the Ritz-values. We will show now that in the reduction to Hessenberg-like form, the part of the matrix already in Hessenberg-like form will have as eigenvalues the rational Ritz-values.

Let us first consider the unitary similarity transformation to Hessenberg form. Assume we have a matrix with its QR -factorization $A = QR$. As proved in Section 3.1 the reduction to Hessenberg form coincides with a transformation of the QR -factorization to the form $H = V^H A V = \hat{Q}\hat{R}$, where \hat{Q} is a unitary Hessenberg and \hat{R} is upper triangular. Assume that the unitary Hessenberg matrix is proper³ and hence the upper triangular matrix $R(1:n-1, 1:n-1)$ is nonsingular (see Section 6.1). In this case we see that the part already in Hessenberg form corresponds to $[\mathbf{e}_1, \dots, \mathbf{e}_k]^T V^H A V [\mathbf{e}_1, \dots, \mathbf{e}_k] = [\mathbf{v}_1, \dots, \mathbf{v}_k]^H A [\mathbf{v}_1, \dots, \mathbf{v}_k]$. The vectors \mathbf{v}_i ($1 \leq i \leq k$) span the Krylov space $\mathcal{K}_k(A, \mathbf{v}_1)$, that is $\mathcal{K}_k(A, \mathbf{v}_1) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$, see, e.g., [29]. Consider now the reduction to Hessenberg-like form. We start with the QR -factorization of the matrix $A = QR$. The reduction to Hessenberg-like form results in a factorization of $V^H A V = \hat{Q}\hat{R}$, where \hat{Q} is now a lower unitary Hessenberg matrix, and \hat{R} is still upper triangular. Assume the unitary matrix to be irreducible and the upper triangular matrix to be nonsingular.

We see that the matrix part $[\mathbf{v}_1, \dots, \mathbf{v}_k]^H A [\mathbf{v}_1, \dots, \mathbf{v}_k]$ already in final Hessenberg-like form has the rational Ritz-values, since $\mathcal{K}_k(A^{-1}, \mathbf{v}_1) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_k\}$ by Theorem 20. As already stressed in Section 6.2, the properness of the Hessenberg-like matrix does not necessarily imply nonsingularity of the matrix R . In case of a singular R , however, computing A^{-1} is impossible and rational Krylov does not exist. Singularity of R does, however, not pose any constraints on the unitary similarity transformation to Hessenberg-like form. The convergence behavior resembles still the convergence behavior of rational Krylov sequences (see Remark 21).

Example 23 In this example we will plot Ritz-value behavior for a symmetric matrix having 100 eigenvalues equally spaced in the interval $[1/n, 1]$. In Figure 1, the horizontal axes denotes the order of the part of the matrix already in Hessenberg or Hessenberg-like form. The vertical axes represent an interval containing the eigen-

³ Otherwise we have a reducible Hessenberg matrix and then we need a suitable vector for reconstructing the corresponding Krylov sequence.

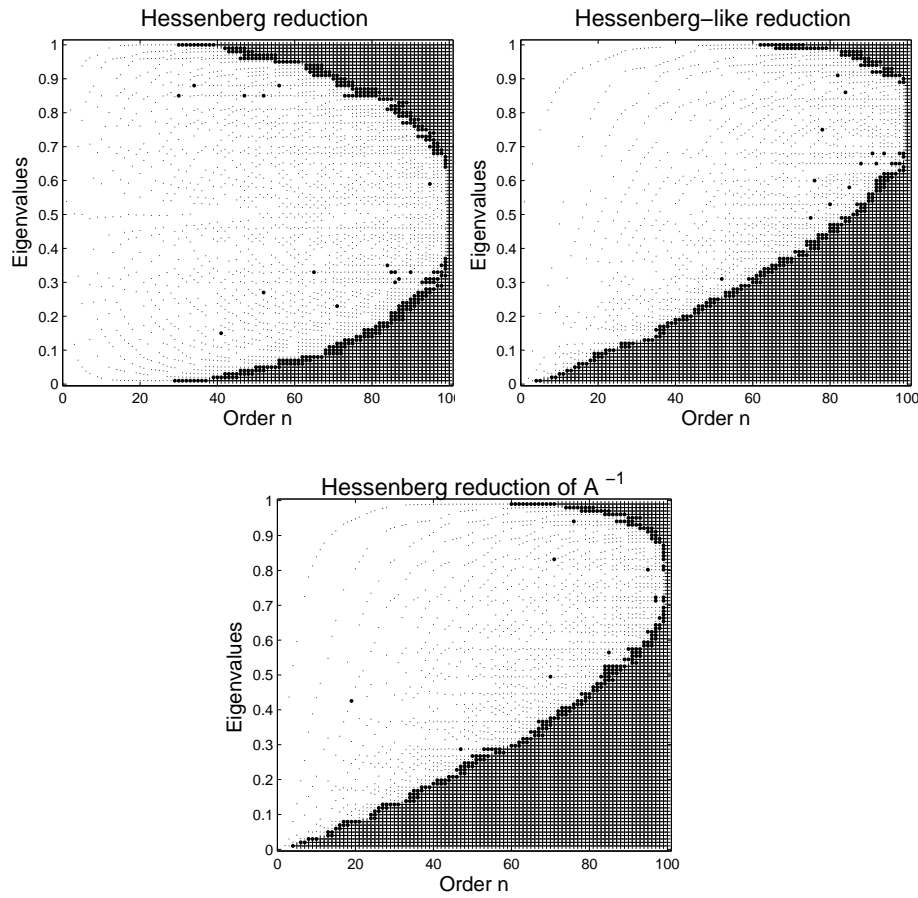


Fig. 1 Convergence plots of Ritz-values. The horizontal axes denotes the order of the part of the matrix already in Hessenberg or Hessenberg-like form. The vertical axes represent an interval containing the eigenvalues of the original matrix. The size and type of dots refers to the accuracy of the approximation. A small dot is depicted if a Ritz-value approximates an eigenvalue in the range $[10^{-2.5}, 10^{-5}]$, a bigger dot if the approximation lies within $[10^{-5}, 10^{-7.5}]$ and a plus sign if the approximation is better than $10^{-7.5}$.

values of the original matrix. In the figures, the size and type of dots refers to the accuracy of the approximation (see the caption of Figure 1).

Three convergence plots of the Ritz-values are shown in Figure 1. The leftmost figure shows the standard Ritz-value convergence behavior, the rightmost the convergence related to the reduction to Hessenberg-like form and one can see that this convergence pattern is almost identical to the pattern received when performing the Hessenberg reduction on A^{-1} , where convergence of the inverse of the eigenvalues is plotted towards the eigenvalues of A .

It can be seen from the figures, that the convergence to small eigenvalues is faster when reducing a matrix to Hessenberg-like form, rather than Hessenberg form, this is a result of the rational Ritz-value behavior.

8 Conclusions and future research

In this paper a new alternative approach was discussed for constructing unitary similarity transformations algorithmically unifying the reductions to Hessenberg-like and Hessenberg form. With respect to the reduction procedure proposed in [23], there are two main differences. First, the new procedure is computationally advantageous and is algorithmically closer to the reduction to Hessenberg form enabling transformations to generalized Hessenberg(-like) form. Second, the new reduction procedure inherits the rational Ritz-value convergence behavior, whereas the classical procedure has the traditional Ritz-value convergence behavior.

It was proven in [13], that Rational Krylov with shifts has a close relation with semiseparable plus diagonal matrices. Unfortunately the semiseparable plus diagonal case does not fit in the novel framework provided here. Future research will focus on this aspect and search for an alternative unitary similarity transformation to semiseparable plus diagonal form, inheriting the rational Krylov behavior, with shifts.

References

1. Beckermann, B., Güttel, S., Vandebril, R.: On the convergence of rational Ritz-values. *SIAM Journal on Matrix Analysis and Applications* **31**(4), 1740–1774 (2010)
2. Bevilacqua, R., Bozzo, E., Del Corso, G.M.: Transformations to rank structures by unitary similarity. *Linear Algebra and its Applications* **402**, 126–134 (2005)
3. Bevilacqua, R., Del Corso, G.M.: Structural properties of matrix unitary reduction to semiseparable form. *Calcolo* **41**(4), 177–202 (2004)
4. Bindel, D., Demmel, J.W., Kahan, W., Marques, O.A.: On computing Givens rotations reliably and efficiently. *ACM Transactions on Mathematical Software* **28**(2), 206–238 (2002)
5. Bini, D., Boito, P., Eidelman, Y., Gemignani, L., Gohberg, I.: A fast implicit QR eigenvalue algorithm for companion matrices. *Linear Algebra and its Applications* **432**(8), 2006–2031 (2010)
6. Bultheel, A., Van Barel, M.: *Linear Algebra, Rational Approximation and Orthogonal Polynomials, Studies in computational mathematics*, vol. 6. North-Holland, Elsevier Science B.V., Amsterdam, Netherlands (1997)
7. Chesnokov, A., Deckers, K., Van Barel, M.: A numerical solution of the constrained weighted energy problem. *Journal of Computational and Applied Mathematics* **235**(4), 950–965 (2010)
8. Delvaux, S., Van Barel, M.: A Givens-weight representation for rank structured matrices. *SIAM Journal on Matrix Analysis and Applications* **29**(4), 1147–1170 (2007)
9. Delvaux, S., Van Barel, M.: A QR-based solver for rank structured matrices. *SIAM Journal on Matrix Analysis and Applications* **30**(2), 464–490 (2008)
10. Delvaux, S., Van Barel, M.: Unitary rank structured matrices. *Journal of Computational and Applied Mathematics* **215**(1), 268–287 (2008)
11. Demmel, J.W.: *Applied Numerical Linear Algebra*. SIAM, Philadelphia, Pennsylvania, USA (1997)
12. Eidelman, Y., Gemignani, L., Gohberg, I.C.: On the fast reduction of a quasiseparable matrix to Hessenberg and tridiagonal forms. *Linear Algebra and its Applications* **420**(1), 86–101 (2007)
13. Fasino, D.: Rational Krylov matrices and QR-steps on Hermitian diagonal-plus-semiseparable matrices. *Numerical Linear Algebra with Applications* **12**(8), 743–754 (2005)
14. Golub, G.H., Van Loan, C.F.: *Matrix Computations*, third edn. Johns Hopkins University Press, Baltimore, Maryland, USA (1996)

15. Kuijlaars, A.B.J.: Which eigenvalues are found by the Lanczos method? *SIAM Journal on Matrix Analysis and Applications* **22**(1), 306–321 (2000)
16. Kuijlaars, A.B.J.: Convergence analysis of Krylov subspace iterations with methods from potential theory. *SIAM Review* **48**(1), 3–40 (2006)
17. Mastronardi, N., Chandrasekaran, S., Van Huffel, S.: Fast and stable reduction of diagonal plus semiseparable matrices to tridiagonal and bidiagonal form. *BIT* **41**(1), 149–157 (2003)
18. Mastronardi, N., Schuermans, M., Van Barel, M., Vandebril, R., Van Huffel, S.: A Lanczos-like reduction of symmetric structured matrices into semiseparable ones. *Calcolo* **42**(3–4), 227–241 (2005)
19. Parlett, B.N.: The Symmetric Eigenvalue Problem, *Classics in Applied Mathematics*, vol. 20. SIAM, Philadelphia, Pennsylvania, USA (1998)
20. Ruhe, A.: Rational Krylov sequence methods for eigenvalue computation. *Linear Algebra and its Applications* **58**, 391–405 (1984)
21. Van Barel, M., Fasino, D., Gemignani, L., Mastronardi, N.: Orthogonal rational functions and diagonal plus semiseparable matrices. In: F.T. Luk (ed.) *Advanced Signal Processing Algorithms, Architectures, and Implementations XII, Proceedings of SPIE, Bellingham, Washington, USA*, vol. 4791, pp. 167–170 (2002)
22. Van Barel, M., Fasino, D., Gemignani, L., Mastronardi, N.: Orthogonal rational functions and structured matrices. *SIAM Journal on Matrix Analysis and Applications* **26**(3), 810–829 (2005)
23. Van Barel, M., Vandebril, R., Mastronardi, N.: An orthogonal similarity reduction of a matrix into semiseparable form. *SIAM Journal on Matrix Analysis and Applications* **27**(1), 176–197 (2005)
24. Van Barel, M., Vandebril, R., Van Dooren, P., Frederix, K.: Implicit double shift *QR*-algorithm for companion matrices. *Numerische Mathematik* **116**(2), 177–212 (2010)
25. Vandebril, R., Van Barel, M., Mastronardi, N.: *Matrix Computations and Semiseparable Matrices, Volume I: Linear Systems*. Johns Hopkins University Press, Baltimore, Maryland, USA (2008)
26. Vandebril, R., Van Barel, M., Mastronardi, N.: *Matrix Computations and Semiseparable Matrices, Volume II: Eigenvalue and Singular Value Methods*. Johns Hopkins University Press, Baltimore, Maryland, USA (2008)
27. Vandebril, R., Van Camp, E., Van Barel, M., Mastronardi, N.: On the convergence properties of the orthogonal similarity transformations to tridiagonal and semiseparable (plus diagonal) form. *Numerische Mathematik* **104**, 205–239 (2006)
28. Vandebril, R., Van Camp, E., Van Barel, M., Mastronardi, N.: Orthogonal similarity transformation of a symmetric matrix into a diagonal-plus-semiseparable one with free choice of the diagonal. *Numerische Mathematik* **102**, 709–726 (2006)
29. Watkins, D.S.: *The Matrix Eigenvalue Problem: GR and Krylov Subspace Methods*. SIAM, Philadelphia, USA (2007)
30. Wilkinson, J.H.: *The Algebraic Eigenvalue Problem*. Numerical Mathematics and Scientific Computation. Oxford University Press, New York, USA (1988)